R E P O R T R E S U M E S

ED 018 200 JC 680 082

TEST BIAS--VALIDITY OF THE SCHOLASTIC APTITUDE TEST FOR NEGRO AND WHITE STUDENTS IN INTEGRATED COLLEGES.

BY- CLEARY, T. ANNE
EDUCATIONAL TESTING SERVICE, PRINCETON, N.J.

REPORT NUMBER ETS-RB-66-31-JUNE-1966 PUB DATE JUN 66
REPORT NUMBER ETS-RDR-65-6-18
EDRS PRICE MF-$0.25 HC-$1.12 26P.

DESCRIPTORS- *JUNIOR COLLEGES, *EDUCATIONAL TESTING, PREDICTIVE ABILITY (TESTING), *APTITUDE TESTS, *CULTURE FREE TESTS, *TEST VALIDITY, HIGHER EDUCATION, TESTING, GROUP TESTS, PROGNOSTIC TESTS, COMPARATIVE TESTING, SCHOLASTIC APTITUDE TEST, COLLEGE ENTRANCE EXAMINATION BOARD,

FOR THIS RESEARCH, A TEST WAS SAID TO BE BIASED FOR MEMBERS OF A SUBGROUP OF THE POPULATION IF, IN THE PREDICTION OF A CRITERION FOR WHICH THE TEST WAS DESIGNED, CONSISTENT NONZERO ERRORS OF PREDICTION ARE MADE FOR MEMBERS OF THE SUBGROUP. SAMPLES OF NEGRO AND WHITE STUDENTS FROM THREE INTEGRATED COLLEGES WERE STUDIED. IN THE TWO EASTERN COLLEGES, NO SIGNIFICANT DIFFERENCES IN THE REGRESSION LINES WERE FOUND. IN ONE COLLEGE IN THE SOUTHWEST, SIGNIFICANT DIFFERENCES WERE FOUND, BUT IT WAS THE NEGRO SCORES THAT WERE OVERPREDICTED. THUS, IN ONE OF THE THREE SCHOOLS, THE SCHOLASTIC APTITUDE TEST WAS FOUND TO BE SLIGHTLY BIASED, BUT IN FAVOR OF THE NEGRO STUDENT. AS THE SCHOOLS USED IN THIS STUDY DO NOT REPRESENT THE FULL SPECTRUM OF COLLEGES IN THE UNITED STATES, NO GENERAL CONCLUSIONS CAN BE REACHED. (AUTHOR/HH)

ERIC

ED018200

COLLEGE ENTRANCE EXAMINATION BOARD
RESEARCH AND DEVELOPMENT REPORTS

•

RESEARCH BULLETIN
RB-66-31 JUNE 1966

TEST BIAS:
Validity of the Scholastic Aptitude Test for Negro and White Students in Integrated Colleges

•

T. Anne Cleary
Developmental Research Division, ETS

FL 680 082

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY
BERKELEY, CALIFORNIA

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

College Entrance Examination Board

RESEARCH AND DEVELOPMENT REPORTS

RDR-65-6, No. 18

TEST BIAS: VALIDITY OF THE SCHOLASTIC APTITUDE TEST FOR NEGRO AND WHITE STUDENTS IN INTEGRATED COLLEGES

T. ANNE CLEARY, Developmental Research Division, ETS

JC 680 082

Research Bulletin

RB-66-31

June 1966

EDUCATIONAL TESTING SERVICE

Princeton, New Jersey

UNIVERSITY OF CALIF.
LOS ANGELES

FEB 23 1968

CLEARINGHOUSE FOR
JUNIOR COLLEGE
INFORMATION

Test Bias: Validity of the Scholastic Aptitude Test for Negro and White Students in Integrated Colleges

Abstract

For this research, a test was said to be biased for members of a subgroup of the population if, in the prediction of a criterion for which the test was designed, consistent nonzero errors of prediction are made for members of the subgroup. Samples of Negro and white students from three integrated colleges were studied. In the two eastern colleges, no significant differences in the regression lines were found. In the one college in the southwest, significant differences were found, but it was the Negro scores which were overpredicted. Thus, in one of the three schools, the Scholastic Aptitude Test was found to be slightly biased, but biased in favor of the Negro student.

Test Bias: Validity of the Scholastic Aptitude Test for Negro and White Students in Integrated Colleges

T. Anne Cleary[1]

In a recent paper, Cleary and Hilton (1966) discussed one possible interpretation of test bias. According to that definition, an item of a test is considered to be biased for members of a particular group if the item produces an uncommon discrepancy between the performance of that group and the performance of other groups. That is, the members of the group obtain an average score which differs from the average score of other groups by more or less than expected from performance on other items of the same test. In terms of the analysis of variance, bias was defined as an item-group interaction. On the basis of their data, Cleary and Hilton concluded that the items in the Preliminary Scholastic Aptitude Test are not biased, and that, if the PSAT is discriminatory, the discrimination is not largely attributable to particular sets of items, but to the test as a whole.

Another definition of bias is possible. The definition of bias used in the present study is concerned with the test as a whole used as a predictor: A test is biased for members of a subgroup of the population if, in the prediction of a criterion for which the test was designed, consistent nonzero errors of prediction are made for members of the subgroup. In other words, the test is biased if too high or too low a criterion score is consistently predicted for members of the subgroup when the common regression line is used. With this definition of bias, there may be a connotation of "unfair," particularly if the use of the test produces a prediction that is too low. The

present research was concerned with this second definition of bias: The prediction of college grade averages from the Scholastic Aptitude Test (SAT) for Negro and white students in integrated colleges was studied.

The validity of the SAT as a predictor of college grades for Negroes in Negro colleges appears to be at least as good as the typical validity for white students. Hills, Klock, and Lewis (1963) report correlations of SAT verbal and mathematics scores with first year average grades for freshmen entering both the Negro and white colleges in the Georgia State University System. The lowest correlations for both male and female students were found in the white colleges rather than in the Negro colleges. The validity of the test in Negro colleges is made more striking by the fact that the standard deviation of the scores in the Negro colleges was approximately half that in the white colleges.

Data reported by Hills (1964) for the four academic years 1959 through 1962 in the Georgia State University System were subjected to analyses of variance by Biaggio and Stanley (1964). They found that, when a correction for restriction in range was applied, the correlations of test scores with freshman grades were significantly higher for the Negroes than for the non-Negroes. When the restriction in range was not considered, they found that the correlations were significantly higher for non-Negro females than for Negro females, but there were no significant differences among males.

Stanley, Biaggio, and Porter (1966) extended the Biaggio-Stanley (1964) study to cover six years, 1959 to 1964. When correlations with grade-point average, corrected for restriction in range in the predominantly Negro colleges and transformed into Fisher's Z, were subjected to four analyses of variance

(SAT-V for men and women, SAT-M for men and women), they were found to be significantly higher in the Negro colleges. When the original correlations were used in the analyses of variance, no significant differences between Negro and non-Negro males were found, but the correlations for non-Negro females were significantly higher than for Negro females. Stanley, Biaggio, and Porter concluded that SAT-type test scores are valid for the prediction of the college grades of Negroes competing with Negroes and taught primarily by Negroes.

McKelpin (1965) studied the prediction of freshman grades from SAT scores and high school average in the predominantly Negro liberal arts college, North Carolina College at Durham. He found validities that were as high as those commonly reported in the literature.

Roberts (1962) found that, in a sample of 129 Fisk freshmen, SAT-V scores had a correlation of .63 with freshman grade-point average, and SAT-M scores, a correlation of .68. In 1964, Roberts reported the correlations for 1962 freshmen in eight Negro colleges with sample sizes ranging from 40 to 203. The median correlations with freshman grade-point average were:

	Male	Female	Total
SAT Verbal	.52	.49	.50
SAT Math	.46	.51	.47

These correlations are similar to those observed in other populations.

When SAT scores have been used in combination with high school rank, similar multiple correlations have been found in both Negro and white colleges (Olsen, 1957; Roberts, 1964).

A question has been raised, however, about the validity of the SAT for predicting academic success of Negro students in integrated colleges. Clark and Plotkin (1963) studied a group of students who had applied for aid from the National Scholarship Service and Fund for Negro Students in order to enter interracial colleges in the years 1952 to 1956. Complete information was not available for the entire sample, and at times it is difficult to determine which subsample was used for a particular comparison. Nevertheless, Clark and Plotkin suggest that perhaps the SAT is not a valid predictor of academic success for Negroes in integrated colleges. They found that while the SAT did discriminate between those who completed college with a B+ or higher average and those who completed college with a C+ or lower average, it did not discriminate between those who completed college and those who did not graduate. A possible explanation for the lack of relationship between grades and SAT scores is severe restriction in range; those students for whom complete information was available were a highly selected group. Campbell (1964) has pointed out that the colleges attended by these students varied widely in degree of selectivity and that it is not unlikely that those of higher ability went to more selective colleges where their grades were perhaps lower than might be expected at a less selective school. Since the same weight was given to grades achieved regardless of college attended, the relationship between grades and ability would be attenuated.

Clark and Plotkin (1963, p. 21) also state that the academic performance of the students they studied was far beyond the level that would be indicated by such predictive indices as College Board scores. Whatever the inadequacies of the Clark and Plotkin study, such a discrepancy deserves further investigation.

Purpose

An important aspect of bias is that concerned with the predictive validity of the test. If the regression of the criterion on the test is the same for different groups, the test cannot be said to be biased in terms of its predictive validity. If the intercepts of the regression lines are different, consistent nonzero errors of prediction will be made within each group. The purpose of this research was to study the regression of college grades on the SAT for Negro and white students in integrated colleges. Because high school rank-in-class is generally used with the SAT for the prediction of grades, rank-in-class was included in the analysis when possible. To determine whether differences in the regressions for Negro and white students were due to differences in curriculum and therefore in the criterion, a sample of white students matched with the Negro students on curriculum was also studied.

Sample

Two major difficulties were encountered in the selection of a sample for this research. In order to compare the regression lines for Negro and white students, it was necessary to find a sufficient number of Negro and white students in the same college. In initial inquiries at various integrated colleges, it was discovered that administrators and faculty tended to overestimate the number of Negro students on the campus. When an actual count was made at some of these colleges, there were too few Negro students to make the analysis feasible. Considering the percentage of the population of the United States that is Negro, the scarcity of Negro students in the integrated colleges is disturbing.

A second difficulty encountered was the identification of the Negro students. Most schools had no record of the race of their individual students.

Perhaps, this situation will change in the future as it is realized that records are necessary for any investigation of bias or, more generally, equal opportunity.

Three schools were used in the study:

School 1 is an eastern, state-supported institution with approximately 5000 male students. The race of the students was identified by having two persons examine independently the standard identification pictures in the school files. Wherever there was disagreement, a third judge was used. If agreement could not be reached, the student was classified as white. Corroboration was obtained from a list of Negro students provided by the NAACP: five students not on the NAACP list had been classified as Negro, and one student on the NAACP list had been classified as white. The five students not on the NAACP list were retained as Negroes after further examination of the identification pictures. The race code of the one student who was on the NAACP list but who had not been classified as Negro was changed to Negro.

Three samples were selected from school 1:

Group 1: All Negro students,

Group 2: A sample of white students matched with the Negro students on curriculum and class, and

Group 3: A random sample of the white students.

School 2 is also an eastern state-subsidized school with approximately 10,000 full-time students. The Negro students were again identified by two persons examining the school identification pictures. When the curriculum and class were tabulated for the Negro students, it was discovered that most of the Negro students (84 out of 148) were freshmen in Liberal Arts. The

remaining Negro students were scattered throughout the other eight curricula and three classes. To make the analysis less complex, only two groups of students were used:

Group 1: All Negro freshmen in Liberal Arts, and

Group 2: A random sample of the white freshmen in Liberal Arts.

School 3 is a state-supported institution in the southwest with approximately 6000 students. Race was identified by the Admissions Office. Three groups were used:

Group 1: All Negro students,

Group 2: A sample of white students matched with the Negro students on sex, class, and curriculum, and

Group 3: A random sample of the white students.

Variables

The criterion in each school was grade-point average (GPA). In all cases the grade-point average was converted to a 1 to 4 scale in which 4 represents the high end of the scale or a grade of A. Unfortunately, the grade-point averages obtained from the different schools have slightly different origins.

In school 1, where all four classes were used, the grade-point average used was the average obtained at the end of one year in the school. In most cases, the average was obtained at the end of the freshman year; in a few cases of transfer students, however, the average was based on more advanced courses. Only freshman Liberal Arts students were used in school 2, so all grade-point averages were from the end of the freshman year. All four classes of school 3 were used and the grade-point average was the latest cumulative average obtained by the student.

The primary predictors were the Scholastic Aptitude Test verbal (SAT-V) and mathematical (SAT-M) scores. In all cases these scores were obtained from the school records.

For students in school 2 and school 3, it was possible to obtain high school rank-in-class (HSR). In school 2, rank was recorded in quintiles, but these were converted so that the resulting scores were: 10, 30, 50, 70, and 90, with 90 representing the first quintile and therefore the top of the class. In school 3, rank was recorded as a normalized score ranging from 25 to 75, with 75 indicating the top of the class. In school 3, rank-in-class was available for only about 50%, so the analyses were repeated using high school grade average (HSA), which was recorded on a scale ranging from 1 to 14, with 14 representing the equivalent of an A+.

Method of Analysis

Correlations were computed among the variables available for each group within each school. For the correlations, all students who had each pair of scores were used for the calculation of the correlations between those scores.

To determine whether the regressions of grades on SAT scores and high school rank were different for the groups of students within each of the three schools, the regression tests of the analysis of covariance were used. The calculations were performed by a method due to Beaton (1964).

The model for the regression tests with two predictors is

$$\hat{Y}_{ig} = \mu + B_1 V_{ig} + B_2 M_{ig} + \mu_g + b_{1g} V_{ig} + b_{2g} M_{ig} ,$$

where

$\hat{Y}_{ig}$ is the predicted criterion score for individual i in group g;

μ is the constant term common to all groups;

B_1 is the component of the regression coefficient for the predictor, V, that is common to all groups;

B_2 is the component of the regression coefficient for the predictor, M, that is common to all groups;

μ_g is the constant term for group g;

b_{1g} and b_{2g} are the regression coefficients that are applied only in group g.

The method of analysis makes it possible to test two hypotheses:

(1) Equality of Slopes: $b_{1g} = b_{2g} = 0$. This hypothesis states that, within each group, the validities of the predictors are the same. If the hypothesis is true and the b's are removed from the model, then the only remaining factor unique to the individual groups is the constant, μ_g. The results of the previous studies of Negro students in Negro colleges indicated that the slopes would be equal within groups. If this hypothesis is rejected, the second test cannot be performed.

(2) Equality of Intercepts (given that the slopes are equal): $\mu_g = 0$ (for all g). μ_g is the constant term that is unique to group g. If all μ_g are not zero, then consistent nonzero errors of prediction are being made within the groups and the test must be considered biased by the definition of this study.

Results

Tables 1, 2, and 3 give the intercorrelations within each of the groups in the three schools. In school 1, the correlations of SAT-V with GPA are almost identical in all three groups. SAT-M, however, has its lowest correlations in

the Negro sample and its highest correlations in the random white sample. It would seem that curriculum differences are, at least in part, contributing to the lower SAT-M correlations in the Negro group, as the SAT-M correlations in the matched white sample are also lower than those in the random white sample. The standard deviation of SAT-M is smaller in the Negro sample than in the other two groups, but this is not an adequate explanation of the reduced correlation because the standard deviation of SAT-V is also smaller in the Negro group.

Insert Tables 1, 2, and 3 about here

In school 2, none of the validities is very impressive. All of the correlations for the Negroes involving high school rank-in-class are essentially zero. High school rank-in-class was a gross measure, quintiles, but this would not account for the differences between the two groups. The near zero correlations among SAT-V, SAT-M, and HSR in the Negro group are perhaps caused by the selection procedure of the college. If a composite score is used for selection and the selection ratio is small for a particular group, reduced correlations among the elements of the composite will be observed in the selected group.

In Table 3, the similarity of the correlations for all three groups is rather striking in view of the discrepancies among the variances and means. All the correlations with GPA are quite high.

The results of the regression tests are presented in Tables 4, 5, 6, and 7. The first hypothesis, equality of slopes, is not rejected in any of the three schools. The second hypothesis, equality of intercepts, is not rejected in school 1 or school 2, but is in school 3.

Insert Tables 4, 5, 6, and 7 about here

At the bottom of Tables 4, 5, 6, and 7 are the equations of within-group regression lines and the common regression line. The similarity of the predictions from each of the equations can be seen by substituting common values of the predictors into each equation. In school 1, if a student has a score of 500 on both SAT-V and SAT-M, his predicted grade-point averages from the different equations will be: Negro, 1.86; matched white, 1.87; random white, 1.95; common line, 1.91. If the student has scores equal to the average scores given in Table 1 for Negroes, his predicted grade-point averages will be: Negro, 1.82; matched white, 1.89; random white, 1.98; common line, 1.92.

In school 2, if a student has SAT scores of 500 and rank-in-class of 50, his predicted grade-point averages will be: Negro, 1.92; white, 1.85; common line, 1.86. If the average scores for Negroes are used, the predicted grade-point averages are: Negro, 1.84, white, 1.87; common line, 1.86. Clearly, in either school, it makes little difference which regression line is used with these predictor scores. The differences between the predicted scores are small and insignificant, but in both schools the common equation predicts a higher score than the Negro equation when predictor scores are equal to the average scores for the Negro students.

In school 3, a significant difference in intercepts was found in both analyses. The differences are not striking, but the large sample size makes the test powerful. If the regression lines from Table 6 are used, a student with SAT scores of 500 and HSR of 50 will have predicted scores: Negro, 2.20;

matched white, 2.58; random white, 2.27; common line, 2.26. If the student has the average scores of Negroes, his predicted scores will be: Negro, 1.96; matched white, 2.46; random white, 2.53; common line, 2.44. For both sets of predictor scores the grade-point average of the Negro students is overpredicted when the common regression line is used. If the regression lines in Table 7 are used, a student with SAT scores of 500 and HSA of 10 will have predicted scores: Negro, 2.34; matched white, 2.68; random white, 2.59; common line, 2.58. If the student has scores equal to the average scores of Negro students, his predicted scores will be: Negro, 1.82; matched white, 2.28; random white, 2.31; common line, 2.23. Again, for both sets of predictor scores, the Negro student's grade-point average will be overpredicted when the common regression line is used.

Conclusions

The schools used in this study do not represent the full spectrum of colleges in the United States, so general conclusions cannot be reached. In the three colleges studied, however, there was little evidence that the Scholastic Aptitude Test was biased as a predictor of college grades. In the two eastern schools, there were no significant differences in the regression lines for Negro and white students. In the one college in the southwest, the regression lines for Negro and white students were significantly different: the Negro students' scores were slightly overpredicted by the use of the common regression line. Thus, where the Scholastic Aptitude Test was found to be biased, it was biased in favor of the Negro student.

References

Beaton, A. E. The use of special matrix operators in statistical calculus. Research Bulletin 64-51. Princeton, N.J.: Educational Testing Service, 1964.

Biaggio, Angela B., & Stanley, J. C. Prediction of freshman grades at southern state colleges. Paper read at the IX Inter-American Congress of Psychology, Miami, Florida, December 1964. Madison, Wisconsin: Laboratory of Experimental Design. 43 Pp. (Mimeographed)

Campbell, J. Testing of culturally different groups. Research Bulletin 64-34. Princeton, N.J.: Educational Testing Service, 1964.

Clark, K. B., & Plotkin, L. The Negro student at integrated colleges. New York: National Scholarship Service and Fund for Negro Students, 1963.

Cleary, T. Anne, & Hilton, T. L. An investigation of item bias. Research Bulletin 66-17. Princeton, N.J.: Educational Testing Service, 1966.

Hills, J. R. Prediction of college grades for all public colleges of a state. Journal of Educational Measurement, 1964, 1, 155-159.

Hills, J. R., Klock, J. C., & Lewis, Sandra. Freshman norms for the University System of Georgia, 1961-1962. Atlanta, Ga.: Office of Testing and Guidance, Regents of the University System of Georgia, 1963.

McKelpin, J. P. Some implications of the intellectual characteristics of freshmen entering a liberal arts college. Journal of Educational Measurement, 1965, 2, 161-166.

Olsen, Marjorie. Summary of main findings on the validity of the CEEB Tests of Developed Ability as predictors of college grades. Statistical Report 57-14. Princeton, N.J.: Educational Testing Service, 1957.

Roberts, S. O. Studies in identification of college potential. Nashville, Tennessee: Department of Psychology, Fisk University, 1962. (Mimeographed)

Roberts, S. O. Comparative validity study of CEEB and CIEP test programs. Nashville, Tennessee: Department of Psychology, Fisk University, 1964. (Mimeographed)

Stanley, J. C., Biaggio, Angela B., & Porter, A. C. Relative predictability of freshman grade-point averages from SAT scores in Negro and white southern colleges. Paper read at the 1966 conventions of the American Educational Research Association and the National Council on Measurement in Education. Madison, Wisconsin: Laboratory of Experimental Design. (Mimeographed)

Footnote

[1]The author is grateful to William Angoff for suggesting the need for this study and to Thomas L. Hilton, Robert L. Linn, and Lenora C. Segal for extensive assistance throughout the project.

Table 1

School 1

Means, Standard Deviations, and Intercorrelations

Group 1 (Negro)

	Mean	Standard Deviation	Intercorrelations[a] SAT-V	SAT-M	GPA
SAT-V	495	67	-	.12 (59)	.47 (59)
SAT-M	525	74	.12 (59)	-	.01 (59)
GPA	1.82	.65	.47 (59)	.01 (59)	-

Group 2 (Matched White)

	Mean	Standard Deviation	Intercorrelations[a] SAT-V	SAT-M	GPA
SAT-V	557	83	-	.14 (60)	.45 (60)
SAT-M	598	79	.14 (60)	-	.25 (60)
GPA	2.23	.67	.45 (60)	.25 (60)	-

Group 3 (Random White)

	Mean	Standard Deviation	Intercorrelations[a] SAT-V	SAT-M	GPA
SAT-V	542	79	-	.47 (118)	.45 (118)
SAT-M	571	83	.47 (118)	-	.41 (118)
GPA	2.18	.58	.45 (118)	.41 (118)	-

[a]Sample size appears in parentheses below the correlation.

Table 2

School 2

Means, Standard Deviations, and Intercorrelations

Group 1 (Negro)

	Mean	Standard Deviation	Intercorrelations[a] SAT-V	SAT-M	HSR	GPA
SAT-V	486	67	-	.09 (83)	-.13 (67)	.26 (83)
SAT-M	468	68	.09 (83)	-	-.15 (67)	.17 (83)
HSR	60	20	-.13 (67)	-.15 (67)	-	.02 (67)
GPA	1.80	.69	.26 (83)	.17 (83)	.02 (67)	-

Group 2 (White)

	Mean	Standard Deviation	Intercorrelations[a] SAT-V	SAT-M	HSR	GPA
SAT-V	502	80	-	.37 (365)	.27 (346)	.38 (365)
SAT-M	517	85	.37 (365)	-	.22 (346)	.30 (365)
HSR	57	22	.27 (346)	.22 (346)	-	.38 (346)
GPA	1.94	.83	.38 (365)	.30 (365)	.38 (346)	-

[a]Sample size appears in parentheses below the correlation.

Table 3

School 3

Means, Standard Deviations, and Intercorrelations

Group 1 (Negro)

	Mean	Standard Deviation	Intercorrelations[a] SAT-V	SAT-M	HSR	HSA	GPA
SAT-V	338	71	-	.51 (131)	.56 (52)	.37 (131)	.47 (125)
SAT-M	371	66	.51 (131)	-	.63 (52)	.35 (131)	.47 (125)
HSR	63	7	.56 (52)	.63 (52)	-	.89 (52)	.61 (48)
HSA	9.9	1.8	.37 (131)	.35 (131)	.89 (52)	-	.51 (125)
GPA	1.81	.56	.47 (125)	.47 (125)	.61 (48)	.51 (125)	-

Group 2 (Matched White)

	Mean	Standard Deviation	Intercorrelations[a] SAT-V	SAT-M	HSR	HSA	GPA
SAT-V	426	96	-	.59 (258)	.48 (144)	.36 (258)	.53 (242)
SAT-M	455	101	.59 (258)	-	.44 (144)	.36 (258)	.41 (242)
HSR	54	9	.48 (144)	.44 (144)	-	.92 (144)	.67 (135)
HSA	8.7	2.3	.36 (258)	.36 (258)	.92 (144)	-	.64 (242)
GPA	2.28	.68	.53 (242)	.41 (242)	.67 (135)	.64 (242)	-

[a]Sample size appears in parentheses below the correlation.

Table 3 (contd.)

Group 3 (Random White)

	Mean	Standard Deviation	Intercorrelations[a] SAT-V	SAT-M	HSR	HSA	GPA
SAT-V	436	100	-	.62 (2325)	.48 (1300)	.45 (2325)	.47 (2181)
SAT-M	461	101	.62 (2325)	-	.45 (1300)	.40 (2325)	.39 (2181)
HSR	55	9	.48 (1300)	.45 (1300)	-	.89 (1300)	.66 (1236)
HSA	9.0	2.3	.45 (2325)	.40 (2325)	.89 (1300)	-	.64 (2181)
GPA	2.38	.70	.47 (2181)	.39 (2181)	.66 (1236)	.64 (2181)	-

[a]Sample size appears in parentheses below the correlation.

Table 4

School 1. Analysis of Covariance

3 Groups, 2 Predictors: SAT-V (V), and SAT-M (M)

The Model:

$$\hat{Y}_{ig} = \mu + B_1 V_{ig} + B_2 M_{ig} + \mu_g + b_{1g} V_{ig} + b_{2g} M_{ig}$$

Tests:

1. Equality of Slopes: $b_{1g} = b_{2g} = 0$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Probability of Larger F
Total about origin	20841496	237			
Under null hypothesis	695009	232			
Due to hypothesis	16636	4	4159.0	1.40	.236
Error	678373	228	2975.3		

2. Equality of Intercepts: $\mu_g = 0$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Probability of Larger F
Total about origin	20841496	237			
Under null hypothesis	704292	234			
Due to hypothesis	9283	2	4641.5	1.55	.215
Error	695009	232	2995.7		

Within-Group Regression Lines:

Group 1 (Negro): $\hat{Y}_1 = -.254 + .00463V - .00041M$

Group 2 (Matched White): $\hat{Y}_2 = -.636 + .00335V + .00166M$

Group 3 (Random White): $\hat{Y}_3 = -.155 + .00243V + .00178M$

Common Regression Line:

$\hat{Y} = -.434 + .00342V + .00126M$

Multiple correlation = .52

Table 5

School 2. Analysis of Covariance

2 Groups, 3 Predictors: SAT-V (V), SAT-M (M), and High School Rank (R)

The Model:

$$\hat{Y}_{ig} = \mu + B_1V_{ig} + B_2M_{ig} + B_3R_{ig} + \mu_g + b_{1g}V_{ig} + b_{2g}M_{ig} + b_{3g}R_{ig}$$

Tests:

1. Equality of Slopes: $b_{1g} = b_{2g} = b_{3g} = 0$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Probability of Larger F
Total about origin	18258638	413			
Under null hypothesis	2046081	408			
Due to hypothesis	21506	3	7168.7	1.43	.232
Error	2024575	405	4998.9		

2. Equality of Intercepts: $\mu_g = 0$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Probability of Larger F
Total about origin	18258638	413			
Under null hypothesis	2046203	409			
Due to hypothesis	122	1	122.0	.024	.876
Error	2046081	408	5014.9		

Within-Group Regression Lines:

Group 1 (Negro): $\hat{Y}_1 = -1.41 + .00481V + .00154M + .00314R$

Group 2 (White): $\hat{Y}_2 = -.741 + .00271V + .00144M + .01034R$

Common Regression Line:

$\hat{Y} = -.831 + .00298V + .00150M + .00897R$

Multiple correlation = .49

Table 6

School 3. Analysis of Covariance

3 Groups, 3 Predictors: SAT-V (V), SAT-M (M), and High School Rank (R)

The Model:

$$\hat{Y}_{ig} = \mu + B_1V_{ig} + B_2M_{ig} + B_3R_{ig} + \mu_g + b_{1g}V_{ig} + b_{2g}M_{ig} + b_{3g}R_{ig}$$

Tests:

1. Equality of Slopes: $b_{1g} = b_{2g} = b_{3g} = 0$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Probability of Larger F
Total about origin	31534459	1419			
Under null hypothesis	3608585	1413			
Due to hypothesis	11190	6	1865.0	.73	.628
Error	3597395	1407	2556.8		

2. Equality of Intercepts: $\mu_g = 0$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Probability of Larger F
Total about origin	31534459	1419			
Under null hypothesis	3717496	1415			
Due to hypothesis	108911	2	54455.5	21.32	<.001
Error	3608585	1413	2553.8		

Within-Group Regression Lines:

Group 1 (Negro): $\hat{Y}_1 = -1.215 + .00247V + .00166M + .02743R$

Group 2 (Matched White): $\hat{Y}_2 = - .742 + .00216V - .00035M + .04143R$

Group 3 (Random White): $\hat{Y}_3 = - .859 + .00146V + .00031M + .04412R$

Common Regression Line:

$\hat{Y} = -.819 + .00178V + .00036M + .04008R$

Multiple correlation = .68

Table 7

School 3. Analysis of Covariance

3 Groups, 3 Predictors: SAT-V (V), SAT-M (M), and High School Average (A)

The Model:

$$\hat{Y}_{ig} = \mu + B_1 V_{ig} + B_2 M_{ig} + B_3 A_{ig} + \mu_g + b_{1g} V_{ig} + b_{2g} M_{ig} + b_{3g} A_{ig}$$

Tests:

1. Equality of Slopes: $b_{1g} = b_{2g} = b_{3g} = 0$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Probability of Larger F
Total about origin	53039632	2548			
Under null hypothesis	6404921	2542			
Due to hypothesis	26435	6	4405.8	1.75	.107
Error	6378486	2536	2515.2		

2. Equality of Intercepts: $\mu_g = 0$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	Probability of Larger F
Total about origin	53039632	2548			
Under null hypothesis	6657499	2544			
Due to hypothesis	252578	2	126289.0	50.12	<.0001
Error	6404921	2542	2519.6		

Within-Group Regression Lines:

Group 1 (Negro): $\hat{Y}_1 = -.572 + .00180V + .00182M + .1104A$

Group 2 (Matched White): $\hat{Y}_2 = -.172 + .00219V + .00032M + .1594A$

Group 3 (Random White): $\hat{Y}_3 = +.110 + .00141V + .00040M + .1571A$

Common Regression Line:

$\hat{Y} = -.0139 + .00174V + .00057M + .1445A$

Multiple correlation = .67